# REGULAR TRANSITION MATRICES

JOSEPH TOBIN

## 1. Introduction

In this paper, we explore properties of regular transition matrices using the guidance of Friedberg, Insel, and Spence [1]. We will start off with basic definitions and cover assumed prior knowledge. Then, we will deduce the relevant theorems and develop a whole swath of properties for regular transition matrices and stationary vectors . Finally, we will explore applications of these theorems.

We rely heavily on results shown in previous parts of section 5.3 in addition to results shown in section 5.1 and 5.2 of [1]. Additionally, previous knowledge of the cosine law and Jordan Canonical Form is assumed, and we use results shown by Ben Pioso in [2].

## 2. Basic Definitions and Notes

**Definition 2.1** (Matrix Limit)**.** Let $L$, $A_1$, $A_2$, ..., be $n \times p$ matrices having complex entries. The sequence $A_1, A_2, \ldots$ is said to **converge** to the $n \times p$ matrix $L$, called the **limit** of the sequence, if $\forall 1 \le i \le n$ and $\forall 1 \le j \le p$

$$lim_{m\to\infty}(A_m)_{ij} = L_{ij}.$$

In this case, we write

$$lim_{m\to\infty}(A_m) = L.$$

**Definition 2.2** (Transition Matrix, Probability Vector)**.** We call an $n \times n$ matrix $M$ with real entries a **transition matrix** if it contains only nonnegative entries and all of its columns sum to 1. We call a column vector $P$ a **probability vector** if it contains only nonnegative entries that sum to 1.

*Remark* 2.3 (Theorem 5.15 [1]).

1. For the rest of the paper, let $u \in \mathbb{C}^n$ be the column vector in which each coordinate equals 1.
2. $M \in M_{n\times n}(\mathbb{R})$ with nonnegative entries is a transition matrix if and only if $M^t u = u$.
3. $v \in \mathbb{R}^n$ with nonnegative coordinates is a probability vector if and only if $u^t v = (1)$.
4. The product of two transition matrices is a transition matrix.
5. The product of a transition matrix and a probability vector is a probability vector.

**Definition 2.4** (Regular)**.** A transition matrix $M$ is called regular if there exists an $s \in \mathbb{N}_{>0}$ such that $M^s$ has only positive entries.

**Definition 2.5** (Row Sum, Column Sum)**.** Let $A \in M_{n\times n}(\mathbb{C})$.

$\forall 1 \le i, j \le n$, define $\rho_i(A)$ to be the sum of the absolute values of the entries of row $i$ of $A$ or

$$\rho_i(A) = \sum_{j=1}^{n} |A_{ij}|$$

and define $\nu_j(A)$ to be the sum of the absolute values of the entries of column $j$ of $A$ or

$$\nu_j(A) = \sum_{i=1}^{n} |A_{ij}|$$

Then the **row sum** of A or $\rho(A)$ and the **column sum** or $\nu(A)$ are defined as:

*Date*: 27 November 2017.

$$\rho(A) = max\{\rho_i(A) \forall 1 \leq i \leq n\}$$

$$\nu(A) = max\{\nu_j(A) \forall 1 \leq j \leq n\}$$

We also use several key theorems presented earlier in the text. We present them without proof below.

**Theorem 1** (Theorem 5.12 [1]). Let $A_1,\ A_2, \ldots$ be a sequence of $n \times p$ matrices with complex entries that converge to the matrix $L$. Then $\forall P \in M_{r\times n}(\mathbb{C}), Q \in M_{p\times s}(\mathbb{C})$, we have

$$lim_{m\to\infty} PA_m = PL$$

and

$$lim_{m\to\infty} A_m Q = LQ$$

**Theorem 2** (Theorem 5.13 [1]). Let $A \in M_{n\times n}(\mathbb{C})$, and for the rest of the paper let $S = \{\lambda \in \mathbb{C} : |\lambda| < 1$ or $\lambda = 1\}$.

Then $lim_{m\to\infty} A^m$ exists if and only if both the following conditions hold:

(1) Every eigenvalue of $A$ is contained in $S$.
(2) If 1 is an eigenvalue of $A$, then the dimension of the eigenspace corresponding to 1 equals the multiplicity of 1 as an eigenvalue of $A$.

**Theorem 3** (Theorem 5.14 [1]). Let $A \in M_{n\times n}(\mathbb{C})$. Then $lim_{m\to\infty} A^m$ exists if the following two conditions hold:

(1) Every eigenvalue of $A$ is contained in $S$.
(2) $A$ is diagonalizable.

**Theorem 4** (Gerschgorin's Disk Theorem Corollary 3 [1]). If $\lambda \in \mathbb{C}$ is an eigenvalue of a transition matrix, then $|\lambda| \leq 1$.

**Theorem 5** (Theorem 5.17 [1]). Every transition matrix has 1 as an eigenvalue.

In the next three lemmas, $F$ is a field and $A \in M_n(F)$.

**Lemma 6** (Exercise 5.1.14 [1]). $A, A^t$ have the same characteristic polynomial and hence the same eigenvalues.

**Lemma 7** (Exercise 5.2.13 [1]). For any eigenvalue $\lambda$ of $A, A^t$, let $E_\lambda, E_\lambda{}'$ denote the corresponding eigenvalues for $A, A^t$ respectively.

Then for any $\lambda$, $dim(E_\lambda)\ =\ dim(E_\lambda{}')$.

**Lemma 8** (Exercise 5.1.15b [1]). Let $x \in F^n$ be an eigenvector of $A$ corresponding to the eigenvalue $\lambda$. For any positive integer $m$, we have the following:

(1) $x$ is an eigenvector of $A^m$ corresponding to the eigenvalue $\lambda^m$.
(2) Let $E_\lambda$ denote the eigenspace of $A$ corresponding to eigenvalue $\lambda$ and $E_\lambda{}'$ the eigenspace of $A^m$ corresponding to $\lambda$. Since $x$ is an eigenvector of $A$ corresponding to the eigenvalue $\lambda$ implies that $x$ is an eigenvector of $A^m$ corresponding to $\lambda^m$, we can conclude $x \in E_\lambda$ implies $x \in E_\lambda{}'$. Thus $E_\lambda \subseteq E_\lambda{}'$.

**Theorem 9** (proven by Ben Pioso in [2]). Let $A \in M_{n\times n}(\mathbb{C})$. Then $lim_{m\to\infty} A^m$ exists if and only if

(1) Every eigenvalue of $A$ is contained in $S$ (recall $S = \{\lambda \in \mathbb{C} : |\lambda| < 1$ or $\lambda = 1\}$).
(2) If 1 is an eigenvalue of $A$, then every Jordan-1-block has size $1 \times 1$.

**Lemma 10.** Let $J_{n_i}(\lambda)$ be the Jordan block of size $n_i \times n_i$ corresponding to eigenvalue $\lambda$ of some matrix $A \in M_{n\times n}(\mathbb{C})$ where $n_i \geq 2$. Then $\forall \ell \in \mathbb{N}_{>0}$ we have

$$(J_{n_i}(\lambda))^\ell_{12} = \ell\lambda^{\ell-1}$$

.

## 3. Theorems

**Lemma 11.** Let $x_1, x_2, \dots x_n \in \mathbb{C}$, and assume $x_1 \neq 0$.

If $|\sum_{i=1}^{n} x_i| = \sum_{i=1}^{n} |x_i|$ then $\forall\ 1 \leq i \leq n\ \exists\ c_i$ such that $x_i = c_i x_1$ and $c_i \in \mathbb{R}_{\geq 0}$ (where $c_1 = 1$).

*Proof.* We will prove this by induction on $n$.

Base Case: $n = 2$ Suppose $|x_1 + x_2| = |x_1| + |x_2|$.

If $x_2$ equals 0, then the property above is trivially satisfied with $c_2 = 0$.

Otherwise we can consider $-x_1, x_2, x_1 + x_2$ as vectors in the complex plane that form a triangle with $|x_1|, |x_2|, |x_1 + x_2|$ as the lengths of the sides of this triangle. But then by the cosine rule $|x_1 + x_2|^2 = |x_1|^2 + |x_2|^2 - 2|x_1||x_2| cos\Theta$, where $\Theta$ is the angle between $-x_1$ and $x_2$. Thus $(|x_1| + |x_2|)^2 = |x_1|^2 + |x_2|^2 + 2|x_1||x_2| = |x_1|^2 + |x_2|^2 - 2|x_1||x_2| cos\Theta$ which implies $1 = -cos\Theta$ and thus $\Theta = \pi$

But if the angle between $-x_1$ and $x_2$ is $\pi$ then the angle between $x_1$ and $x_2$ is 0. Therefore since $x_1, x_2 \neq 0$, we can conclude $x_2 = c_2 x_1$, where $c_2 > 0$, $c_2 \in \mathbb{R}$.

Inductive step: Assume that $|\sum_{i=1}^{n-1} x_i| = \sum_{i=1}^{n-1} |x_i|$ implies $\forall\ 1 \leq i \leq n-1\ \exists\ c_i \geq 0$ such that $x_i = c_i x_1$ and $c_i \in \mathbb{R}$(where $c_1 = 1$). Additionally, assume $|\sum_{i=1}^{n} x_i| = \sum_{i=1}^{n} |x_i|$.

First, we want to be able to show that $|\sum_{i=1}^{n-1} x_i| = \sum_{i=1}^{n-1} |x_i|$ so we can use the inductive hypothesis. By the triangle inequality we have

$$|\sum_{i=1}^{n} x_i| \leq |\sum_{i=1}^{n-1} x_i| + |x_n|$$

and thus

$$\sum_{i=1}^{n} |x_i| = |\sum_{i=1}^{n} x_i| \leq |\sum_{i=1}^{n-1} x_i| + |x_n|$$

which implies

$$\sum_{i=1}^{n-1} |x_i| \leq |\sum_{i=1}^{n-1} x_i|.$$

And by the triangle inequality again for the first $n-1$ terms we have:

$$|\sum_{i=1}^{n-1} x_i| \leq \sum_{i=1}^{n-1} |x_i|$$

which combined implies

$$\sum_{i=1}^{n-1} |x_i| = |\sum_{i=1}^{n-1} x_i|$$

and thus by assumption $\forall\ 1 \leq i \leq n-1\ \exists\ c_i$ such that $x_i = c_i x_1$ and $c_i \in \mathbb{R}_{\geq 0}$ (where $c_1 = 1, x_1 \neq 0$).

Therefore

$$|\sum_{i=1}^{n} x_i| = |(\sum_{i=1}^{n-1} c_i x_1) + x_n| = |(\sum_{i=1}^{n-1} c_i) x_1 + x_n|$$

and

$$\sum_{i=1}^{n} |x_i| = \sum_{i=1}^{n-1} |x_i| + |x_n| = |\sum_{i=1}^{n-1} x_i| + |x_n| = |(\sum_{i=1}^{n-1} c_i) x_1| + |x_n|$$

.

But we know by the base case that $|(\sum_{i=1}^{n-1} c_i) x_1| + |x_n| = |(\sum_{i=1}^{n-1} c_i) x_1 + x_n|$ implies that $\exists k \in \mathbb{R}_{\geq 0}$ such that $x_n = k((\sum_{i=1}^{n-1} c_i) x_1)$. Let $c_n = k(\sum_{i=1}^{n-1} c_i)$. Since $c_1, c_2 \dots c_{n-1}, k \geq 0$ and are elements of $\mathbb{R}$, we know $c_n \in \mathbb{R}_{\geq 0}$.

Therefore $\exists c_n \in \mathbb{R}_{\geq 0}$ such that $x_n = c_n x_1$, and thus we have shown $\forall\ 1 \leq i \leq n\ \exists\ c_i \in \mathbb{R}_{\geq 0}$ such that $x_i = c_i x_1$ (where $c_1 = 1$).

□

**Theorem 12** (Theorem 5.18 pg. 298 [1])**.** Let $A \in M_{n\times n}(\mathbb{R})$ be a matrix in which each entry is positive and let $\lambda \in \mathbb{C}$ be an eigenvalue of $A$ such that $|\lambda| = \rho(A)$. Then $\lambda = \rho(A)$ and $\{u\}$ is a basis for $E_\lambda$, where $u \in \mathbb{C}^n$ is the column vector in which each coordinate equals 1.

*Proof.* First we want to show that $\{u\}$ is a basis for $E_\lambda$, considered as a subspace of $\mathbb{C}^n$. To show this, we are going to use the fact that $|\lambda| = \rho(A)$ to derive several equalities giving us information about $A$.

Let $v \in \mathbb{C}^n$ be an eigenvector of $A$ corresponding to $\lambda$ with coordinates $v_1, v_2, \ldots, v_n$. Now choose $k$ such that $v_k$ is the coordinate of $v$ with the largest absolute value and let $b = |v_k|$.

Then

$$|\lambda|\ b = \ |\lambda|\ |v_k| = |\lambda v_k|$$

But if $\lambda$ is an eigenvalue of $A$, then $Av = \lambda v$ and thus $\forall\ 1 \le i \le n$, $\lambda v_i = \sum_{j=1}^n A_{ij} v_j$. Thus

$$|\lambda v_k| = |\sum_{j=1}^n A_{kj} v_j|$$

By the triangle inequality and the absolute value multiplication rules,

$$|\sum_{j=1}^n A_{kj} v_j| \le \sum_{j=1}^n |A_{kj} v_j| = \sum_{j=1}^n |A_{kj}||v_j|$$

Since we know $b = |v_k| \ge |v_j|\ \forall\ 1 \le j \le n$ and similarly $\rho(A) \ge \rho_i(A)\ \forall\ 1 \le i \le n$, we know

$$\sum_{j=1}^n |A_{kj}||v_j| \le \sum_{j=1}^n |A_{kj}|b = b\sum_{j=1}^n |A_{kj}| = b\rho_k(A) = \rho_k(A)b \le \rho(A)b$$

But since we know $|\lambda| = \rho(A)$, we know the three inequalities used above are actually equalities.

(1)

$$|\sum_{j=1}^n A_{kj} v_j| = \sum_{j=1}^n |A_{kj} v_j|$$

(2)

$$\sum_{j=1}^n |A_{kj}||v_j| = \sum_{j=1}^n |A_{kj}|b$$

(3)

$$\rho_k(A)b = \rho(A)b$$

But now we can use this to show that $\{u\}$ is a basis for $E_\lambda$. By lemma 11, we know (1) above holds if and only if $A_{kj}v_j$ are nonnegative multiples of each other. Without loss of generality, we can assume they are multiples of a complex number $z$ with $|z| = 1$ by adjusting the constants $c_1 \ldots c_n$. Then $\exists c_1, c_2, \ldots c_n \in \mathbb{R}_{\ge 0}$ such that $A_{kj}v_j = c_j z$. Since $A_{kj} > 0$, we can say $v_j = \frac{c_j z}{A_{kj}}$.

By (2) above, using again $A_{kj} > 0$, we know $\forall 1 \le j \le n$ , $|v_j| = b$ and therefore

$$b = |v_j| = |\frac{c_j z}{A_{kj}}|$$

But $A_{kj} > 0$ by assumption and $c_j$ is nonnegative. Thus

$$b = |\frac{c_j z}{A_{kj}}| = |\frac{c_j}{A_{kj}}||z| = |\frac{c_j}{A_{kj}}| * 1 = \frac{c_j}{A_{kj}}\ \forall 1 \le j \le n$$

Since $v_j = \frac{c_j z}{A_{kj}}$, this gives us $v_j = bz\ \forall j$ and thus

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} bz \\ bz \\ \vdots \\ bz \end{bmatrix} = bzu \tag{3.1}$$

Thus any eigenvector $v$ of $A$ corresponding to $\lambda$ can be expressed as a scalar multiple of $u$, and thus $\{u\}$ is a basis for $E_\lambda$, and $u$ is an eigenvector of $A$.

Now note that because $A$ has all positive values and $u$ has all positive values, then $Au$ has all positive values. Therefore, because $\lambda u = Au$, $\lambda u$ has all positive values. Since $u$ has each coordinate as 1, we can conclude $\lambda > 0$ and thus $\lambda = |\lambda| = \rho(A)$.

□

*Corollary* 12.1 (Corollary 1 pg. 299 [1] ). Let $A \in M_{n\times n}(\mathbb{R})$ be a matrix in which each entry is positive and let $\lambda \in \mathbb{C}$ be an eigenvalue of $A$ such that $|\lambda| = \nu(A)$. Then $\lambda = \nu(A)$ and $E_\lambda$ has dimension 1.

*Proof.* Consider $A^t$ and let $E_\lambda,\ E_\lambda{}'$ be the eigenspaces of $\lambda$ corresponding to $A,\ A^t$ respectively. We know by lemma 6 that $A$ and $A^t$ have the same eigenvalues. Thus $A^t$ is a matrix in which each entry is positive and has eigenvalue $\lambda$.

Now notice that because the columns of $A$ are the rows of $A^t$ we know $\nu(A) = \rho(A^t)$ and thus $|\lambda| = \rho(A^t)$. Thus $A^t$ is a matrix with all positive entries and with an eigenvalue $\lambda$ satisfying $|\lambda| = \rho(A^t)$, and by Theorem 12, the basis of $E_\lambda{}' = \{u\}$ and $\lambda = \rho(A^t) = \nu(A)$.

Then $dim(E_\lambda{}') = 1$ and by lemma 7 , we know $dim(E_\lambda) = dim(E_\lambda{}') = 1$.

Thus, we have shown $\lambda = \nu(A)$ and $dim(E_\lambda) = 1$.

□

*Corollary* 12.2 (Corollary 2 pg. 299 [1]). Let $A \in M_{n\times n}(\mathbb{R})$ be a transition matrix in which each entry is positive, and let $\lambda \in \mathbb{C}$ be an eigenvalue of $A$ such that $\lambda \neq 1$. Then $|\lambda| < 1$, and the eigenspace corresponding to the eigenvalue 1 has dimension 1.

*Proof.* We know by Theorem 4 that if $\lambda$ is an eigenvalue of a transition matrix, then $|\lambda| \leq 1$. Thus if $|\lambda| \neq 1$, then $|\lambda| < 1$. Suppose $|\lambda| = 1$. We know by definition of a transition matrix that the sum of each column of $A$ is 1. In other words $\nu_i(A) = 1$ for all $1 \leq i \leq n$. Hence $\nu(A) = 1$, and thus $|\lambda| = \nu(A) = 1$. But this means by corollary 12.1 that $\lambda = \nu(A) = 1$. By contraposition $\lambda \neq 1$ implies $|\lambda| \neq 1$ and thus $\lambda \neq 1$ implies $|\lambda| < 1$ .

Since $A$ is a transition matrix, then by Theorem 5 we know that 1 is an eigenvalue of $A$. Therefore we have an all-positive-entry matrix $A$ with eigenvalue $1 = \nu(A)$, and thus by the previous corollary, we know $dim(E_1) = 1$.

□

**Theorem 13** (Theorem 5.19 pg. 298)**.** Let $A$ be a regular transition matrix and let $\lambda$ be an eigenvalue of $A$. Then

(1) $|\lambda| \leq 1$
(2) If $|\lambda| = 1$, then $\lambda = 1$ and $dim(E_\lambda) = 1$.

*Proof.* We know (1) by Theorem 4.

Since $A$ is regular, we know by definition $\exists s \in \mathbb{N}_{>0}$ such that $A^s$ has only positive entries. Moreover $A^s$ and $A^{s+1}$ are transition matrices because $A$ is a transition matrix and the product of transition matrices is a transition matrix. We now split this proof into parts:

(1) Because $A$ is a transition matrix and the entries of $A^s$ are positive, we know the entries of $A^{s+1} = A^s(A)$ are positive. More specifically,

$$A^{s+1}_{ij} = \sum_{k=1}^{n} A^s_{ik} A_{kj}$$

and thus because for any given $i, j$, $\forall\ 1 \leq k \leq n\ A_{kj} \geq 0$ and $A^s_{ik} > 0$ and $\exists\ k$ such that $A_{kj} > 0$ (since the column sums to 1), we can conclude $A^{s+1}_{ij} > 0$.

(2) Suppose $|\lambda| = 1$, then we know by lemma 8 that if $\lambda$ is an eigenvalue of $A$, then $\lambda^s$, $\lambda^{s+1}$ are eigenvalues of $A^s$, $A^{s+1}$ respectively. Because $|\lambda| = 1$, we know $|\lambda^s| = |\lambda^{s+1}| = |\lambda| = 1$.

(3) Because each entry of $A^s$, $A^{s+1}$ is positive and both matrices are transition matrices, we know by corollary 12.2 that for any eigenvalues $\lambda*$ of $A^s$, $A^{s+1}$ such that $\lambda* \neq 1$, then $|\lambda*| < 1$. Thus because $|\lambda^s| = |\lambda^{s+1}| = 1$, we can conclude $\lambda^s = \lambda^{s+1} = 1$ and therefore $\lambda = 1$.

(4) Let $E_\lambda$ and $E_\lambda{}'$ be the eigenspaces of $A$, $A^{s+1}$ respectively corresponding to $\lambda = 1$. Then by lemma 8 $E_\lambda \subseteq E_\lambda{}'$. But then by corollary 12.2 we know $dim(E_\lambda{}') = 1$, and thus $dim(E_\lambda) = 1$.

□

*Corollary* 13.1 (Corollary Pg. 300). Let $A$ be a regular transition matrix that is diagonalizable. Then $\lim_{m\to\infty} A^m$ exists.

*Proof.* We know by Theorem 3 that if $A \in M_{n\times n}(\mathbb{C})$ is diagonalizable and has every eigenvalue contained in $S$, then $\lim_{m\to\infty} A^m$ exists.

Thus because $A$ is diagonalizable by assumption, we just need to show for all eigenvalues $\lambda$ of $A$, $\lambda \in S$. But we know by Theorem 13 that $\forall$ eigenvalues $\lambda$, $\lambda = 1$ or $|\lambda| < 1$ and thus $\lambda \in S$. Thus we can conclude that $\lim_{m\to\infty} A^m$ exists.

□

The following lemmas use Jordan Canonical Form to prove Theorem 16.

**Definition 3.1.** For any $A \in M_{n\times n}(\mathbb{C})$, define the norm of $A$ by

$$||A|| = max\{|A_{ij}|\ \forall\ i, j\}$$

**Lemma 14** (Exercise 20 of Section 7.2)**.** Let $A, B \in M_{n\times n}(\mathbb{C})$. Then $||AB|| \leq n||A||||B||$.

*Proof.* By definition $||AB|| = max\{|(AB)_{ij}|\ \forall\ i, j\}$. Next, we know that $|(AB)_{ij}| = |\sum_{k=1}^n A_{ik}B_{kj}|$ and $\forall\ i, j, k\ |A_{ik}| \leq ||A||, |B_{kj}| \leq ||B||$. Thus all together

$$||AB|| = max\{|(AB)_{ij}|\ \forall\ i, j\} = max\{|\sum_{k=1}^n A_{ik}B_{kj}|\ \forall\ i, j\} \leq \sum_{k=1}^n ||A||||B||| = n||A||||B||$$

Thus we can conclude $||AB|| \leq n||A||||B||$.

□

**Lemma 15** (Modified Exercise 21 of Section 7.2 [1])**.** Let $A \in M_{n\times n}(\mathbb{R})$ be a transition matrix. Since $\mathbb{C}$ is an algebraically closed field, $A$ has a Jordan canonical form $J$ to which $A$ is similar in $M_{n\times n}(\mathbb{C})$. Let $P \in M_{n\times n}(\mathbb{C})$ be an invertible matrix such that $P^{-1}AP = J$. Then we have the following:

(1) $||A^m|| \leq 1$ for every positive integer $m$.
(2) $\exists c > 0$ such that $||J^m|| \leq c$ for every positive integer $m$.
(3) Each Jordan block of $J$ corresponding to the eigenvalue $\lambda = 1$ is a $1 \times 1$ matrix.

*Proof.*

(1) If $A$ is a transition matrix, then every column sums to 1 and every $A_{ij} \geq 0$. Thus $\forall\ i, j\ 0 \leq A_{ij} \leq 1$ and thus $max\{|A_{ij}|\ \forall\ i, j\} = ||A|| \leq 1$. The same applies to $A^m$ for any $m \in \mathbb{N}$ since $A^m$ is also a transition matrix.

(2) If $J = P^{-1}AP$, then $J^m = P^{-1}A^mP$ and $||J^m|| = ||P^{-1}A^mP||$. By lemma 14 we have that $||P^{-1}A^mP|| \leq n||P^{-1}||||A^mP|| \leq n^2||P^{-1}||||A^m||||P||$. But we know by (1) that $||A^m|| \leq 1$. Thus

$$|J^m|| \leq ||P^{-1}A^mP|| \leq n^2||P^{-1}||||A^m||||P|| \leq n^2||P^{-1}|||P||$$

But since $P, P^{-1}$ are fixed, $||P||, ||P^{-1}||$ are constant and thus we have a constant $c = n^2||P^{-1}||||P||$ such that $||J^m|| \leq c\ \forall m \in \mathbb{Z}_{>0}$

(3) By way of contradiction, suppose there exists a Jordan block of $J_{n_i}(\lambda)$ in $J$ corresponding to $\lambda = 1$ of size $n_i \times n_i$ where $n_i \neq 1$ (which implies $n_i \geq 2$). But we know by lemma 10 for $n_i \geq 2$ that $(J_{n_i}(\lambda))^m_{12} = m\lambda^{m-1} = m * 1 = m$. But $(J_{n_i}(\lambda))^m_{12}$ is an element of $J^m$ and thus $||J^m||$ is not bounded, a contradiction!

Thus $J_{n_i}(\lambda) = J_1(1)$ is a $1 \times 1$ matrix.

□

**Theorem 16** (Theorem 5.20 [1])**.** Let $A$ be an $n \times n$ regular transition matrix. Then:

(1) $\lim_{m\to\infty} A^m$ exists.
(2) $L = \lim_{m\to\infty} A^m$ is a transition matrix.
(3) $LA = AL = L$.
(4) The columns of $L$ are identical and equal to the unique probability vector $v$ that is an eigenvector corresponding to the eigenvalue 1.
(5) For any probability vector $w$, $\lim_{m\to\infty} A^m w = v$.

*Proof.*

(1) By lemma 15, since $A \in M_{n\times n}(\mathbb{R})$ is transition matrix, we know $A$ has a Jordan Canonical Form $J$ to which $A$ is similar such that each Jordan block of $J$ corresponding to the eigenvalue $\lambda = 1$ is a $1 \times 1$ matrix.

Additionally, by Theorem 13, we know all eigenvalues $\lambda$ of $A$ are contained in $S$.

Thus, by Theorem 9, we know that $\lim_{m\to\infty} A^m$ exists.

(2) By Remark 2.3 , we know $L$ is a transition matrix if and only if $L^t u = u$, where $u$ is the column vector where every entry is equal to 1. But due to transposition rules, we can equivalently say $u^t L = u^t$. But

$$u^t L = u^t \lim_{m\to\infty} A^m = \lim_{m\to\infty} u^t A^m$$

But $A^m$ is a transition matrix which means $\forall m \; u^t A^m = u^t$ and thus

$$\lim_{m\to\infty} u^t A^m = \lim_{m\to\infty} u^t = u^t$$

Thus $L$ is a transition matrix.

(3) By Theorem 1, we know $AL = \lim_{m\to\infty} AA^m = \lim_{m\to\infty} A^{m+1} = \lim_{m\to\infty} A^m A = LA$.

But $\lim_{m\to\infty} A^{m+1} = \lim_{m\to\infty} A^m = L$.

Thus $LA = AL = L$.

(4) We know $AL = L$ by item 3. Let $L_i$ be the $ith$ column vector of $L$. Then because $AL = L$, we know $AL_i = L_i \forall 1 \leq i \leq n$. Thus $L_i$ is an eigenvector of $A$ corresponding to eigenvalue $\lambda = 1$. Additionally, because $L$ is a transition matrix by item 2 we know $L_i$ is a probability vector.

By way of contradiction assume there exists two distinct eigenvectors $v_1, v_2$ corresponding to eigenvalue $\lambda = 1$ that are probability vectors. By Theorem 13, we know that $dim(E_\lambda) = 1$ and thus $v_1$, $v_2$ must be multiples of eachother. Thus $\exists c \neq 1 \in \mathbb{C}$ such that $v_2 = cv_1$. But if the sum of the coordinates of $v_1$ is equal to 1, then the sum of the coordinates of $v_2$ is equal to $c * 1 \neq 1$ and $v_2$ is not a probability vector, a contradiction.

Thus there is a unique eigenvector $v$ corresponding to $\lambda = 1$ that is a probability vector, and $L_i = v \; \forall 1 \leq i \leq n$.

(5) Let $w$ be any probability vector and $y = \lim_{m\to\infty} A^m w = Lw$. We want to show $y = v$.

If $y = Lw$, then by item 2 and Remark 2.3 , we know $y$ is a probability vector. Additionally,

$$Ay = A(Lw) = (AL)w = Lw$$

by item 3. But $Lw = y$ and thus $Ay = y$.

Therefore $y$ is a probability vector and an eigenvector of $A$ corresponding to $\lambda = 1$. But $v$ is the unique probability vector which is an eigenvector of $A$ corresponding to $\lambda = 1$, and thus $y = v$. Thus $\lim_{m\to\infty} A^m w = v$.

□

# 4. Applications

**Definition 4.1.** The vector $v$ in Theorem 16(5) is called the **fixed probability vector** or **stationary vector** of the regular transition matrix $A$.

The following is an example 4 of section 5.3 from [1].

> A survey in Persia showed that on a particular day 50% of the Persians preferred a loaf of bread, 30% preferred a jug of wine, and 20% preferred "thou beside me in the wilderness." A subsequent survey 1 month later yielded the following data: Of those who preferred a loaf of bread on the first survey, 40% continued to prefer a loaf of bread, 10% now preferred a jug of wine, and 50% preferred "thou"; of those who preferred a jug of wine on the first survey, 20% now preferred a loaf of bread, 70% continued to prefer a jug of wine, and 10% now preferred "thou"; of those who preferred "thou" on the first survey, 20% now preferred a loaf of bread, 20% now preferred a jug of wine, and 60% continued to prefer "thou."
>
> Assuming that this trend continues, the situation described in the preceding paragraph is a three-state Markov chain in which the slates are the three possible preferences. We can predict the percentage of Persians in each state for each month following the original survey.

Letting the first, second, and third states be preferences for bread, wine, and "thou", respectively, we see that the probability vector that gives the initial probability of being in each state is

$$P = \begin{bmatrix} 0.50 \\ 0.30 \\ 0.20 \end{bmatrix} \tag{4.1}$$

and the transition matrix is

$$A = \begin{bmatrix} 0.40 & 0.20 & 0.20 \\ 0.10 & 0.70 & 0.20 \\ 0.50 & 0.10 & 0.60 \end{bmatrix} \tag{4.2}$$

The probabilities of being in each state $m$ months after the original survey are the coordinates of the vector $A^m P$.

For example, after $m = 1$ month we have

$$AP = \begin{bmatrix} 0.30 \\ 0.30 \\ 0.40 \end{bmatrix} \tag{4.3}$$

which means there's a 30% chance of a randomly selected person having a preference for bread, 30% chance for wine, and 40% chance for "thou".

We can similarly calculate for $m = 2, 3, 4, \ldots$

Since A is regular, the long-range prediction concerning the Persians' preferences can be found by computing the fixed probability vector for $A$. This vector is the unique probability vector $v$ such that $(A - I)v = 0$

If we let

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{4.4}$$

we can solve the following system of linear equations

$$\begin{bmatrix} -0.60v_1 + 0.20v_2 + 0.20v_3 = 0 \\ 0.10v_1 - 0.30v_2 + 0.20v_3 = 0 \\ 0.50v_1 + 0.10v_2 - 0.40v_3 = 0 \end{bmatrix} \tag{4.5}$$

to get

$$\begin{bmatrix} 5 \\ 7 \\ 8 \end{bmatrix} \tag{4.6}$$

as a basis for the solution space of the system. Thus the unique fixed probability vector for $A$ is

$$\begin{bmatrix} \frac{5}{5+7+8} \\ \frac{7}{5+7+8} \\ \frac{8}{5+7+8} \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.35 \\ 0.4 \end{bmatrix}$$

Thus, in the long run. 25% of the Persians prefer a loaf of bread, 35% prefer a jug of wine, and 40% prefer "thou beside me in the wilderness."

We can computationally confirm this by finding $lim_{m\to\infty} A^m$.

Let

$$Q = \begin{bmatrix} 5 & 0 & -3 \\ 7 & -1 & -1 \\ 8 & 1 & 4 \end{bmatrix} \tag{4.7}$$

then

$$Q^{-1}AQ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.2 \end{bmatrix} \tag{4.8}$$

So

$$lim_{m\to\infty} A^m = lim_{m\to\infty}(QQ^{-1}AQQ^{-1})^m = Q(lim_{m\to\infty}(Q^{-1}AQ)^m)Q^{-1}$$

$$= Q\left[lim_{m\to\infty} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}^m\right] Q^{-1} = Q \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^{-1} = \begin{bmatrix} 0.25 & 0.25 & 0.25 \\ 0.35 & 0.35 & 0.35 \\ 0.4 & 0.4 & 0.4 \end{bmatrix}$$

which is as expected.

## References


[1] Friedberg, S.H. and Insel, A.J. and Spence, L.E. *Linear Algebra.* Featured Titles for Linear Algebra (Advanced) Series. Pearson Education, 2003.

[2] Pioso, Ben. *Matrix Limits, Including Jordan Canonical Form.* 2017.